

DESCRIPTIONTHREE-DIMENSIONAL STRUCTURE OF DIPEPTIDYL PEPTIDASE IV5 TECHNICAL FIELD

The present invention relates to a crystal and a three-dimensional structural coordinate of a dipeptidyl peptidase IV, and an application thereof. More specifically, the present invention relates to a crystal and a three-dimensional structural coordinate, a method for obtaining a three-dimensional structural coordinate of a homolog protein of a dipeptidyl peptidase IV, a method for obtaining a three-dimensional structural coordinate of a crystal of a complex of a dipeptidyl peptidase IV with an effector (e.g. inhibitor) of the dipeptidyl peptidase IV, a method for identifying a pharmacophore of an effector (e.g. inhibitor) of for the dipeptidyl peptidase IV, a method for identifying sites affecting the activity of the dipeptidyl peptidase IV, a method for designing, identifying, evaluating or searching an effector (e.g. inhibitor) of the dipeptidyl peptidase IV, and a program and a medium therefor for use of the three-dimensional structural coordinate, which are useful in the development of an effector (e.g. inhibitor) of the dipeptidyl peptidase IV, useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like; and an effector (e.g. inhibitor) of the dipeptidyl peptidase IV useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like.

## BACKGROUND ART

Dipeptidyl peptidase IV (hereinafter also referred to as DPPIV) is a cell membrane protein, which has been found in epithelial cell of small intestine, prostate gland, renal tubule, biliary tract and the like, activated T-cell, B-cell, NK-cell and the like. In the DPPIV, deduced active sites of DPPIV in the C-terminal side are located in extracellular portions and those in the N-terminal side are located in cytoplasm in a living body. Also, there has been suggested the relationship of the above-mentioned DPPIV with the activities of various cytokines such as interleukin-1 $\beta$ , interleukin-2, interleukin-3, interleukin-5, interleukin-6, interleukin-13, tumor necrosis factor- $\beta$  and the like, and activities of various chemokines such as RANTES and the like in immune system [*Rinsho Menneki (Clinical Immunology)*, 34, Revised and Enlarged Edition 19, 45-53, published by Kagaku Hyoronsha (2000), and the like].

As to the dipeptidyl peptidase IV, it has been shown that some amino acid residues can be involved in exhibition of the activity of the dipeptidyl peptidase IV by experiments such as biochemical experiments using inhibitors, experiments using mutants produced by site-directed mutagenesis [for example, see Misumi et al, *Biochim. Biophys. Acta*, 1131, 333-336 (1992), Ogata et al, *Biochemistry*, 31, 2582-2587 (1992) and the like].

However, it is difficult to know the three-dimensional structures for active sites from the information. Therefore, it is presently difficult to obtain the three-dimensional structural information for identifying, searching, evaluating or designing an interaction of the dipeptidyl peptidase IV and a compound that acts with the dipeptidyl peptidase IV on the level of three-dimensional structure and a

novel compound capable of binding with and acting on the dipeptidyl peptidase IV.

#### DISCLOSURE OF INVENTION

5           A first object of the present invention is to provide a crystal of a dipeptidyl peptidase IV, which is useful for providing a three-dimensional structural coordinate as the information for designing, identifying, evaluating or searching an effector (e.g. inhibitor) of the dipeptidyl peptidase IV useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent  
10 for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like. A second object of the present invention is to provide a three-dimensional structural coordinate of the crystal, which can provide the information for designing, identifying, evaluating or searching an effector (e.g. inhibitor) of the dipeptidyl peptidase IV useful as a  
15 modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like. A third object of the present invention is to provide a method for obtaining a three-dimensional structural coordinate of a homolog protein of the dipeptidyl peptidase IV, whereby  
20 refinement of a three-dimensional structural coordinate of a homolog protein of the dipeptidyl peptidase IV can be more readily performed. Furthermore, a fourth object of the present invention is to provide a method for obtaining a three-dimensional structural coordinate of a crystal of a complex of a dipeptidyl peptidase IV and an effector (e.g. inhibitor) of the dipeptidyl peptidase IV, which  
25 can provide the information for designing, identifying, evaluating or searching an

effector (e.g. inhibitor) of the dipeptidyl peptidase IV which is useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like, and is excellent in avidity, biological activity, biological stability, absorbency to a living body, and which can favorably act on the dipeptidyl peptidase IV. A fifth object of the present invention is to provide a method for identifying a pharmacophore of the dipeptidyl peptidase IV and the effector (e.g. inhibitor) of the dipeptidyl peptidase IV, which can provide the information for designing, identifying, evaluating or searching an effector (e.g. inhibitor) of the dipeptidyl peptidase IV useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like, and is excellent in avidity, biological activity, biological stability, absorbency in a living body, and which can be favorably act on the dipeptidyl peptidase IV. A sixth object of the present invention is to provide a method for designing, identifying, evaluating or searching the effector (e.g. inhibitor) of the dipeptidyl peptidase IV, which can logically and conveniently provide the effector (e.g. inhibitor) of the dipeptidyl peptidase IV useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like, and is excellent in avidity, biological activity, biological stability, absorbency in a living body (*in vivo*), and which can be favorably act on the dipeptidyl peptidase IV. A seventh object of the present invention is to provide the effector (e.g. inhibitor) of the dipeptidyl peptidase IV useful as a modulatory agent of immune

response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like. An eighth object of the present invention is to provide a program and a medium therefor, which can rapidly and conveniently perform design, identification, evaluation or search of the effector (e.g. inhibitor) of the dipeptidyl peptidase IV.

Concretely, the present invention relates to:

- [1] a crystal of a dipeptidyl peptidase IV, having characteristics sufficient to ensure a resolution capable of analyzing its three-dimensional structure up to the side chain level by X-ray crystallographic structural analysis;
- [2] the crystal according to the above [1], wherein the dipeptidyl peptidase IV is a soluble polypeptide comprising a region located at extramembrane in a full-length human dipeptidyl peptidase IV;
- [3] the crystal according to the above [1] or [2], wherein the dipeptidyl peptidase IV is a polypeptide having an amino acid sequence in which a transmembrane region is deleted from the amino acid sequence of SEQ ID NO: 2, and a tag peptide is optionally added to a C-terminal side or N-terminal side thereof;
- [4] the crystal according to any one of the above [1] to [3], wherein the crystal has a space group of  $P2_12_12_1$ , and a lattice constant of the unit cell of  $|a| = 118.0 \pm 5.0 \text{ \AA}$ ,  $|b| = 125.9 \pm 5.0 \text{ \AA}$ ,  $|c| = 136.8 \pm 5.0 \text{ \AA}$ , and  $\alpha = \beta = \gamma = 90^\circ$ , and is orthorhombic;
- [5] the crystal according to any one of the above [1] to [4], wherein the crystal has the structural coordinate shown in Figure 4;
- [6] the crystal according to any one of the above [1] to [4], wherein the

crystal has a structural coordinate different from the structural coordinate as shown in Figure 4 via fluctuation of a protein;

[7] a three-dimensional structural coordinate of a dipeptidyl peptidase IV, comprising the structural coordinate shown in Figure 4;

5 [8] a three-dimensional structural coordinate of a dipeptidyl peptidase IV, comprising a structural coordinate different from the structural coordinate as shown in Figure 4 via fluctuation of a protein;

[9] the three-dimensional structural coordinate according to the above [8], wherein the fluctuation of a protein is a state that is caused by molecular  
10 oscillation or temperature, and exhibits an activity for the dipeptidyl peptidase IV in a living body;

[10] the three-dimensional structural coordinate according to any one of the above [7] to [9], wherein the dipeptidyl peptidase IV is a soluble polypeptide comprising a region located at extramembrane in a full-length human dipeptidyl  
15 peptidase IV;

[11] the three-dimensional structural coordinate according to any one of the above [7] to [10], wherein the dipeptidyl peptidase IV is a polypeptide having an amino acid sequence in which a transmembrane region is deleted from the amino acid sequence of SEQ ID NO: 2, and a tag peptide is optionally added of to a  
20 C-terminal side or N-terminal side thereof;

[12] a three-dimensional structural coordinate of a region in a dipeptidyl peptidase IV, comprising the three-dimensional structural coordinate of the region selected from the group consisting of the following (a) to (d):

(a) a region characterized by Ser 630, Asp 708 and His 740 in the amino acid  
25 sequence of SEQ ID NO: 2, and

all or a part of a group of the amino acid residues located in the adjacent area of each of the Ser 630, Asp 708 and His 740 in the structural coordinate shown in Figure 4 or the three-dimensional structure model defined by the structural coordinate;

- 5 (b) a region characterized by Ser 630, Asp 708 and His 740 in the amino acid sequence of SEQ ID NO: 2, and

all or a part of a group of the amino acid residues comprising amino acids capable of maintaining physicochemical characteristics physiologically equivalent to each of amino acids in the group of the amino acid residues located in the adjacent area of each of Ser 630, Asp 708 and His 740, in the structural coordinate shown in Figure 4 or the three-dimensional structure model defined by the structural coordinate,

- 10 (c) a region characterized by a group of amino acid residues comprising amino acids capable of maintaining physicochemical characteristics physiologically equivalent to each of Ser 630, Asp 708 and His 740 in the amino acid sequence of SEQ ID NO: 2, and

all or a part of a group of the amino acid residues located in the adjacent area of said group of the amino acid residues in the structural coordinate shown in Figure 4 or the three-dimensional structure model defined by the structural coordinate; and

- 20 (d) a region characterized by a group of amino acid residues comprising amino acids capable of maintaining physicochemical characteristics physiologically equivalent to each of Ser 630, Asp 708 and His 740 in the amino acid sequence of SEQ ID NO: 2, and

25 all or a part of a group of amino acid residues comprising amino acids

capable of maintaining physicochemical characteristics physiologically equivalent to each of the amino acids in the group of the amino acid residues located in the adjacent area of said group of the amino acids, in the structural coordinate shown in Figure 4 or the three-dimensional

5 structure model defined by the structural coordinate,

wherein the region in the dipeptidyl peptidase IV is a region involved in binding or interaction between the dipeptidyl peptidase IV and an effector of the dipeptidyl peptidase IV;

10 [13] the three-dimensional coordinate according to the above [12], wherein the physicochemical characteristic is selected from the group consisting of features in shape of a three-dimensional structure, hydrophobicity, electric charge and pK;

15 [14] a method for obtaining a three-dimensional coordinate of a homolog protein of a dipeptidyl peptidase IV, characterized in refining an electron density map of the homolog protein of the dipeptidyl peptidase IV comprising the amino acid sequence of SEQ ID NO: 2, based on all and/or a part of the three-dimensional coordinate of any one of the above [7] to [13], to give a three-dimensional structural coordinate;

20 [15] a method for obtaining a three-dimensional structural coordinate of a crystal of a complex of a dipeptidyl peptidase IV and an effector of the dipeptidyl peptidase IV characterized in using all and/or a part of the three-dimensional structural coordinate of any one of the above [7] to [13], to give a three-dimensional structural coordinate;

25 [16] a method for identifying pharmacophore of an effector of the dipeptidyl peptidase IV, characterized in identifying the pharmacophore based on all and/or



a part of the three-dimensional structural coordinate of any one of the above [7] to [13], and the steric conformation of the effector of the dipeptidyl peptidase IV; [17] a method for designing, identifying, evaluating or searching an effector of a dipeptidyl peptidase IV, characterized in designing, identifying, evaluating or searching a compound capable of acting on the dipeptidyl peptidase IV, based on all and/or a part of the three-dimensional structural coordinate of any one of the above [7] to [13];

[18] the method according to the above [17], wherein the method for designing, identifying, evaluating or searching an effector comprises the steps of:

- (i) identifying a region to be targeted for binding or interaction with the effector in a dipeptidyl peptidase IV, based on all and/or a part of the three-dimensional structural coordinate according to any one of the above [7] to [13] and the steric conformation of the effector of the dipeptidyl peptidase IV;
- (ii) identifying atoms or atomic groups capable of generating in the above region at least one intermolecular interaction selected from the group consisting of covalent bond, ionic interaction, ion-dipole interaction, dipole-dipole interaction, hydrogen bonding, van der Waals force, electrostatic interaction and hydrophobic interaction, with the atoms or atomic groups existing in a candidate compound; and
- (iii) designing a compound based on the information of the above step (i) and/or (ii);

[19] the method according to the above [18], wherein the method further comprises the steps of:

detecting an interaction between the dipeptidyl peptidase IV and the

designed, identified, evaluated or searched candidate compound,  
wherein when an interaction is detected, the candidate compound is identified as  
a compound capable of binding to the dipeptidyl peptidase IV, based on a degree  
of the interaction as an index;

5 [20] the method according to the above [18] or [19], wherein the method  
further comprises the steps of:

contacting the dipeptidyl peptidase IV with the designed, identified,  
evaluated or searched candidate compound and measuring the activity of the  
dipeptidyl peptidase IV,

10 wherein when an activity increases or decreases, the designed, identified,  
evaluated or searched candidate compound is identified as a compound having  
enhancing action or inhibitory action on the activity of the dipeptidyl peptidase  
IV, based on a degree of the increase or decrease as an index;

[21] an effector of the dipeptidyl peptidase IV obtainable by the method of any  
15 one of the above [17] to [20];

[22] a program and a medium therefor for use of the three-dimensional  
structural coordinate of any one of the above [7] to [13], wherein all and/or a part  
of the three-dimensional structural coordinate of any one of the above [7] to [13]  
is recorded;

20 [23] the program and the medium according to the above [22], comprising a  
means for identifying, searching, evaluating or designing a compound capable of  
binding to the dipeptidyl peptidase IV or a compound having an enhancing  
action or inhibitory action on the activity for the dipeptidyl peptidase IV; and

[24] the program and the medium according to the above [23], further  
25 comprising a means for displaying a three-dimensional graphic display of a

molecule.

### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a photomicrograph of a crystal of a dipeptidyl peptidase IV,  
5 wherein the field of view is 4000  $\mu\text{m}$   $\times$  3000  $\mu\text{m}$ .

Figure 2 is a photograph for X-ray diffraction pattern of a crystal of dipeptidyl peptidase IV.

Figure 3 is a photograph showing a three-dimensional structure of a crystal of a dipeptidyl peptidase IV displayed by the program QUANTA  
10 (Accelrys, Inc.).

Figure 4 is a drawing showing a three-dimensional coordinate of a crystal of a dipeptidyl peptidase IV.

### BEST MODE FOR CARRYING OUT THE INVENTION

15 In the present specification, amino acid residues are expressed by using the following abbreviations, which have been adopted by the IUPAC-IUB Commission on Biochemical Nomenclature (CBN). Also, unless explicitly otherwise indicated, the amino acid sequences of peptides and proteins are identified from N-terminal to C-terminal, left terminal to right terminal, the  
20 N-terminal being identified as a first residue. Ala: alanine residue; Asp: aspartate residue; Glu: glutamate residue; Phe: phenylalanine residue; Gly: glycine residue; His: histidine residue; Ile: isoleucine residue; Lys: lysine residue; Leu: leucine residue; Met: methionine residue; Asn: asparagine residue; Pro: proline residue; Gln: glutamine residue; Arg: arginine residue; Ser: serine  
25 residue; Thr: threonine residue; Val: valine residue; Trp: tryptophane residue;

Tyr: tyrosine residue; Cys: cysteine residue.

The crystal of the present invention is a crystal of a dipeptidyl peptidase IV, having a characteristic sufficient to ensure a resolution capable of analyzing its three-dimensional structure up to the side chain level by X-ray crystallographic structural analysis.

The "characteristic sufficient to ensure a resolution capable of analyzing three-dimensional structure up to the side chain level" is, for example,

(1) being in a state that a molecule in a unit cell of a crystal has repeats with high regularity, namely, providing diffraction at high resolution;

(2) having suitable form and size; it is desired that for example, a crystal has at least one side grown to about 0.2 to about 0.5 mm, preferably a cubic crystal having three sides that have similarly grown, or a needle-shaped crystal having a width or thickness of about 0.2 mm or more;

(3) having chemical stability, dynamic stability and physical stability; and the like. In a case of the dipeptidyl peptidase IV, which is a polypeptide having a relatively large molecular weight, the term means characteristics sufficient to ensure a resolution of 3Å or less, preferably 2.8Å or less, more preferably 2.6Å or less.

The dipeptidyl peptidase IV used for the preparation of the crystal of the present invention may have a high purity sufficient for forming the crystal. In the present invention, the dipeptidyl peptidase IV used for the preparation of the crystal includes a soluble polypeptide consisting of a region located at extramembrane in a full-length human dipeptidyl peptidase IV, for example, a polypeptide in which a transmembrane region in the N-terminal side [namely the region including the transmembrane sites (the region containing at least the

amino acid nos: 1-28 of SEQ ID NO: 2, preferably the region of the amino acid nos: 1-32)] is deleted from the amino acid sequence of a full-length human dipeptidyl peptidase IV of SEQ ID NO: 2, and a tag peptide is optionally added to a C-terminal side or N-terminal side of the amino acid sequence. Concrete examples include (I) a polypeptide in which a transmembrane region in the N-terminal side is deleted from the amino acid sequence of a full-length human dipeptidyl peptidase IV of SEQ ID NO: 2; and (II) a polypeptide in which a tag peptide is added to a C-terminal side or N-terminal side of the polypeptide of the above (I). In the polypeptide, since the transmembrane site is deleted therefrom, the polypeptide has excellent characteristics that anchoring to the membrane can be prevented, and the polypeptide is a secretory type and soluble. The tag peptide is not particularly limited. For example, a polyhistidine peptide (an oligopeptide consisting of 4 to 20 of histidine residues) or the like can be preferably used as the tag peptide.

SEQ ID NO: 2 shows the amino acid sequence of a full-length dipeptidyl peptidase IV of human colon.

The full-length dipeptidyl peptidase IV means a polypeptide of a dipeptidyl peptidase IV containing a region comprising a transmembrane site in the N-terminal side. The full-length dipeptidyl peptidase IV includes a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, without being limited thereto, and encompasses its naturally occurring variant, artificially modified variant, a homolog and an ortholog derived from heterogeneous organism, and the like.

Concretely, the full-length dipeptidyl peptidase IV, besides the polypeptide comprising the amino acid sequence of SEQ ID NO: 2, includes

conservative substitution variants, naturally occurring allelic variants and the like. Also, the full-length dipeptidyl peptidase IV includes a polypeptide having at least one, namely one or more conservative amino acid substitutions, as compared to the polypeptide comprising the amino acid sequence of SEQ ID NO: 2.

The polypeptide as described above may be a polypeptide having biological activities (namely dipeptidyl peptidase IV activity) similar to the polypeptide comprising the amino acid sequence of SEQ ID NO: 2. Concretely, there are included, for instance, a polypeptide having homology of usually about 80% or more, preferably about 90% or more, more preferably about 95% or more on the amino acid level, as compared to the full-length amino acid sequence of SEQ ID NO: 2; a polypeptide encoded by a nucleic acid capable of hybridizing with a nucleic acid consisting of the nucleotide sequence of SEQ ID NO: 1 (nucleotide sequence encoding a full-length dipeptidyl peptidase IV of human colon), under stringent conditions, or a complement thereof; and a polypeptide having deletion, substitution or addition of at least one amino acid, namely one or plural amino acids, preferably one or several amino acids in the amino acid sequence of SEQ ID NO: 2.

The number of deletion, substitution or addition of the amino acids may be to an extent that the biological activities [namely, dipeptidyl peptidase IV activity] are not lost, usually in the number of 1 to about 150, preferably 1 to about 75, more preferably 1 to about 40.

The crystallization is carried out by making a solution containing the desired protein (referred to as a protein solution) supersaturated state, based on the characteristics that the protein in solution state converts to non-soluble state

and precipitates as a crystal when specific conditions are satisfied. Concretely, the protein can be precipitated by the following procedures 1. or 2.:

1. elevating the effective concentration of the protein:

concretely, adding a precipitant such as a salt, polyethylene glycol or an organic solvent to a protein solution; reducing an amount of a solvent in the protein solution by evaporation or the like; or the like.

2. reducing a repulsive force, or increasing an attractive force between protein molecules:

concretely, adding an organic solvent such as an alcohol to a protein solution; changing a hydrogen ion concentration (pH) or temperature of the protein solution; or the like.

As the conditions for the crystallization, physical and chemical factors such as a hydrogen ion concentration (pH), a kind of buffer used and a concentration thereof, a kind of a precipitant added and a concentration thereof, protein concentration, salt concentration, temperature and the like can be involved. A method for controlling and investigating the factors includes batch methods, dialysis methods, vapor diffusion methods (hanging-drop method, sitting-drop method and the like) and the like, described, for instance, in Blundell, T. L. et al., *PROTEIN CRYSTALLOGRAPHY*, 59-82 (1976), published by Academic Press, or the like.

The method for crystallization includes the batch methods, dialysis methods, vapor diffusion methods and the like. By the above method, physical and chemical factors such as a hydrogen ion concentration (pH), a kind and a concentration of the buffer used, and a kind and a concentration of the precipitant used, and physical and chemical factors such as protein concentration, salt

concentration and temperature can be also determined.

The hydrogen ion concentration (pH) can be adjusted with a buffer. It is desired that the buffer is a buffer having buffering action in a broad range of pH, and being capable of suppressing precipitation of a non-proteinous crystal  
5 between the co-existing ion in the solution used during crystallization and the precipitant or the like. The buffer includes Tris-hydrochloric acid buffer, phosphate buffer, cacodylate buffer, acetate buffer, citrate buffer, glycine buffer and the like.

The precipitant may be a substance capable of elevating an effective  
10 concentration of the protein or changing a hydrogen ion concentration (pH) of the protein solution. Generally, the precipitant includes salts such as ammonium sulfate, sodium sulfate, sodium phosphate, potassium phosphate, sodium citrate, ammonium citrate, sodium chloride, potassium chloride and ammonium chloride; polyethylene glycols having various average molecular weights of  
15 about 200, about 1000, about 2000, about 4000, about 6000, about 8000, about 20000 or the like; organic solvents such as 2-methyl-2,4-pentadiol, methanol, ethanol, isopropanol, butanol and acetone, and the like.

The protein concentration may be a concentration suitable for crystallization, and it is desired that the protein concentration is, for example, 1  
20 to 50 mg/ml, preferably 5 to 20 mg/ml, more preferably 7 to 15 mg/ml.

It is desired that the temperature conditions are 3° to 25°C, preferably 12° to 22°C.

In the case where the crystallization is carried out by the batch method, the crystallization can be carried out by gradually adding a precipitant solution  
25 comprising a precipitant, buffer and the like, so as to form a layer on the top



layer of the solution containing the dipeptidyl peptidase IV to give a mixture, or by gradually adding the solution comprising the dipeptidyl peptidase IV, so that the solution is an upper layer of the precipitant solution to give a mixture. Here, the mixture is allowed to stand in a tightly closed vessel.

5           In the case where the crystallization is carried out by the dialysis method, the crystallization can be carried out by placing a solution comprising dipeptidyl peptidase IV in a size exclusion semi-permeable membrane, and placing a precipitant solution outside of the size exclusion semi-permeable membrane as a reservoir solution, thereby diffusing the reservoir solution to the solution  
10           comprising the dipeptidyl peptidase IV via the semi-permeable membrane.

          In the case where the crystallization is carried out by the hanging-drop method in the vapor diffusion method, the crystallization can be carried out by placing a mixed solution of a solution comprising the dipeptidyl peptidase IV and a precipitant solution in a closed vessel allowing to be hanged at a position  
15           above the upper space of a reservoir in which the precipitant solution is contained as a reservoir solution, wherein the vapor pressure of the reservoir solution in the reservoir is set to be lower than that of the mixed solution.

          In the case where the crystallization is carried out by the sitting-drop method in the vapor diffusion method, the crystallization can be carried out by  
20           placing a mixed solution comprising a solution comprising the dipeptidyl peptidase IV and a precipitant solution in a closed vessel at a position higher than the liquid surface of a reservoir in which the precipitant solution is contained as a reservoir solution, wherein the vapor pressure of the reservoir solution in the reservoir is set to be lower than that of the mixed solution.

25           The crystallization can be carried out by the sitting-drop method from the

viewpoint of obtaining excellent-quality and large crystal.

When the obtained crystal is a crystal insufficient to ensure the X-ray structural analysis, the crystal may be grown by a seeding method such as macro-seeding method or micro-seeding method, using the obtained crystal as a seed crystal.

When the macro-seeding method is performed, it is desired that the seed crystal is a single crystal that can be isolated by procedures under microscope wherein the seed crystal has excellent external form (having excellent crystallinity). Also, it is desired that the seed crystal is washed with a drop of a solution obtained by diluting the precipitant, for example, by 0.5 to 1.0-fold. It is desired that the solution used for seeding of the seed crystal is a protein solution having a degree of supersaturation that the crystal grows but the crystal nuclei do not grow. On the other hand, when the micro-seeding method is performed, the form and size of the seed crystal are not particularly limited.

The sequence information for the dipeptidyl peptidase IV and cDNA encoding the dipeptidyl peptidase IV can be obtained from a known information source [GenBank/EMBL accession No: X60708; Misumi et al., *Biochim. Biophys. Acta*, **1131**, 333-336, (1992); GenBank/EMBL accession No: M80536; Darmoul et al., *J. Biol. Chem.*, **267**, 4824-4833, (1992)]. Therefore, the dipeptidyl peptidase IV or a soluble polypeptide thereof can be produced by using conventional means for gene engineering on the basis of the above sequence information.

The nucleic acid used for production of the dipeptidyl peptidase IV or a soluble polypeptide thereof may be any nucleic acid in which the encoded polypeptide exhibits a dipeptidyl peptidase IV activity. For example, a nucleic

acid encoding a polypeptide consisting of the amino acid sequence in which a transmembrane region in the N-terminal side (a region containing at least the amino acid nos: 1-28, preferably the region of the amino acid nos: 1-32) is deleted from the full-length human dipeptidyl peptidase IV, and a tag peptide is optionally added to a C-terminal side or N-terminal side of the amino acid sequence.

The nucleic acid can be obtained by, for instance, obtaining a fragment comprising a nucleic acid encoding a full-length dipeptidyl peptidase IV or a part thereof by means of conventional DNA recombination technique, and appropriately arranging the obtained fragment.

SEQ ID NO: 1 shows a sequence of a nucleic acid encoding a full-length dipeptidyl peptidase IV of human colon.

The nucleic acid (DNA or RNA) encoding a full-length dipeptidyl peptidase IV includes, for instance, a nucleic acid comprising human nucleic acids comprising the nucleotide sequence of SEQ ID NO: 1 without being limited thereto, and includes its naturally occurring variant, artificially modified variant, a homolog or ortholog derived from heterogeneous organism.

In other words, besides the nucleic acid comprising the nucleotide sequence of SEQ ID NO: 1, the nucleic acid includes a nucleic acid capable of hybridizing with a nucleic acid comprising the nucleotide sequence of SEQ ID NO: 1 under stringent conditions, more preferably under high-stringent conditions), or a complement thereof (nucleic acid having a complementary sequence).

Concrete examples of the nucleic acid described above include, for instance, a nucleic acid having usually about 70% or more, preferably about 80%

or more, more preferably about 85% or more, still more preferably about 90% or more, still more preferably about 95% or more, homology to the nucleotide sequence of SEQ ID NO: 1, and it is preferable that the polypeptide encoded by the above nucleic acid has a dipeptidyl peptidase IV activity.

5           The dipeptidyl peptidase IV activity can be measured by, for example, incubating in a 1.5 ml reaction mixture [composition: 1.5 mM substrate (Gly-Pro-paranitroanilide), 71 mM glycine-NaOH (pH 8.7)] at 37°C for 10 minutes, and determining the liberated paranitroanilide at the absorbance of 405 nm. One unit (1 U) of a dipeptidyl peptidase IV is defined as an amount of  
10           the enzyme required for liberating 1  $\mu$ mol of paranitroanilide per 1 minute.

          In the present invention, the hybridization under stringent conditions can be carried out as normal stringent conditions by performing hybridization in a hybridization solution having a salt concentration of  $6 \times \text{SSC}$  or an equivalent concentration thereto, under the temperature conditions of 50° to 70°C for about  
15           16 hours, and optionally performing pre-washing with a solution having a salt concentration of  $6 \times \text{SSC}$  or an equivalent concentration thereto, and thereafter performing washing with a solution having a salt concentration of  $1 \times \text{SSC}$  or an equivalent concentration thereof. Furthermore, as the conditions having still  
20           higher stringency (high-stringent conditions), the hybridization can be carried out by washing with a solution having a salt concentration of  $0.1 \times \text{SSC}$  or an equivalent concentration thereto in the above method.

          The dipeptidyl peptidase IV used for the crystallization has purity that can form a crystal, and the purity can be confirmed by conventional means of confirming purity (for example, a method comprising electrophoresing a fraction  
25           by polyacrylamide gel electrophoresis, SDS-polyacrylamide gel electrophoresis

or the like, and visualizing the fraction by silver staining, or the like).

The X-ray structural analysis data of the crystal can be obtained by subjecting the crystal of the present invention to an X-ray crystallographic structural analysis known to one of ordinary skill in the art [for example, see  
5 Blundell, T. L. et al., *PROTEIN CRYSTALLOGRAPHY*, 59-82 (1976), published by Academic Press, and the like], whereby a three-dimensional structural coordinate (a value showing the relationship of the spatial positions of each atom) and a three-dimensional structure model for the crystal can be obtained. Concretely, the three-dimensional structural coordinate of the dipeptidyl  
10 peptidase IV is obtained as an atomic coordinate by procedures comprising the steps of 1) irradiating the crystal of the present invention with a monochromatic X-ray to give an X-ray diffraction pattern, 2) obtaining X-ray diffraction intensity data from the X-ray diffraction pattern, 3) obtaining an electron density map by Fourier transform, and 4) allocating a polypeptide chain and side chain  
15 thereof on the electron density map based on the amino acid sequence of the polypeptide used for the crystal. Furthermore, the three-dimensional structure is clarified by molecule-modeling based on the three-dimensional structural coordinate. Therefore, the three-dimensional structural coordinate of the dipeptidyl peptidase IV obtained from the crystal of the present invention is also  
20 encompassed within the scope of the present invention.

The crystallographic parameters for the crystal are obtained from the X-ray diffraction intensity data of the crystal of the present invention. The crystal of the present invention is an orthorhombic crystal having a space group of  $P2_12_12_1$ , and a lattice constant of the unit cell of  $|a| = 118.0 \pm 5.0\text{\AA}$ ,  
25  $|b| = 125.9 \pm 5.0\text{\AA}$ ,  $|c| = 136.8 \pm 5.0\text{\AA}$ , and  $\alpha = \beta = \gamma = 90^\circ$ . The crystal has a

2.6Å resolution by X-ray crystallographic structural analysis, that is, the crystal has characteristics sufficient to ensure a resolution capable of analyzing up to the side chain level of the polypeptide.

It is a known fact to one of ordinary skill in the art that the same protein can be crystallized even under different conditions. Therefore, the present invention is not limited to only the conditions for crystallization, and the crystal that shows substantially the same crystallographic constants as those in the present invention are also encompassed within the scope of the present invention.

More concretely, the crystal of the dipeptidyl peptidase IV of the present invention has a structural coordinate as shown in Figure 4, or a structural coordinate different from the structural coordinate as shown in Figure 4 via fluctuation of a protein.

The crystal according to the present invention can also be used as a seed crystal for carrying out the crystallization of a polypeptide having a three-dimensional structure similar to that of the dipeptidyl peptidase IV used for, for example, carrying out the crystallization of the dipeptidyl peptidase IV, dipeptidyl peptidase IV-like proteins, homolog proteins and the like, which are derived from other organism species.

When the crystal of the present invention is irradiated with X-ray, a low-temperature measurement may be carried out, as described in Examples set forth below.

The X-ray structural analysis data are converted to a structure factor by evaluating the intensity of X-ray diffraction using MOSFILM Program Package (Version 6.1). Also, in order to obtain the information for the phase, multiple isomorphous replacement method or the like can be performed, for example, as

described in Examples.

In the structural analysis, CCP4 (Collaborative Computational Project, Number 4, 1994, "*The CCP4 Suite: Programs for Protein Crystallography*," *Acta Cryst.* D50, 760-763) program or the like is used.

5           The three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention can be obtained, for example, as follows. Firstly, Fourier transform calculation is carried out using the differences between the diffraction intensity obtained from two kinds of isomorphous replacement crystals of mercury and the diffraction intensity obtained from native crystal, and  
10           investigating the large peaks provided by the heavy atoms (mercury) on the Patterson's diagram to determine the locations of each mercury atoms in the unit cell of the real space. The phase of the crystal structure factor for the native crystal is determined using the obtained location coordinate for the mercury atoms. Furthermore, refinement is performed using the crystal structure factor of  
15           the native crystal and two kinds of the crystal structure factors of the isomorphous replacement crystals of mercury, and the coordinate for each of the mercury atoms is more accurately determined. An electron density map for the crystal of the dipeptidyl peptidase IV in the real space is obtained using the phase of the crystal structure factor of the native crystal calculated from the refined  
20           mercury atoms coordinate. Furthermore, the electron density map is improved by performing smoothing and histogram matching for the electron density map of the solvent region, whereby an electron density map necessary and sufficient for building a molecular model can be obtained. Next, the sites corresponding to the amino acid residues of the dipeptidyl peptidase IV on the electron density  
25           map are identified using QUANTA (manufactured by Accelrys, Inc.) to build the

molecular model to give a three-dimensional structural coordinate.

The three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention is shown in Figure 4. Figure 4 shows the obtained three-dimensional structural coordinates, according to the format of the Protein Data Bank, which is a notation generally used by one of ordinary skill in the art.

The three-dimensional structural coordinates shown in Figure 4 are those represented using the origin of the unit cell of the crystal as the origin of the three-dimensional space. The R factor that is considered as an index for the accuracy of the obtained molecular model is 24.89%, and the free R factor is 30.15%. In addition, the deviation in the interatomic bond distance from the ideal state of the three-dimensional structure (rms-deviation) and the deviation in the bond angle are 0.006Å and 1.305°, respectively. In the case, for instance, the three-dimensional structural coordinate of the present invention is used for the calculation by a computer, a novel structural coordinate obtained as a result of the operation for mathematical transfer, such as translation, rotation, or symmetry in the three-dimensional space without changing the relative configuration of the atoms, is also encompassed within the scope of the present invention. Furthermore, not only all of the three-dimensional structural coordinate of the present invention but also a part thereof are also encompassed within the scope of the present invention.

The three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention can be used, for example, as shown in Figure 3, for three-dimensional graphic displaying of the stereogram of the three-dimensional structure model, and for evaluation of the structure-activity relationship and the quantitative structure-activity relationship. Also, the structural features of the



crystal of the present invention can be more concretely shown using the three-dimensional structural coordinate shown in Figure 4. The evaluation of the structure-activity relationship or quantitative structure-activity relationship by the three-dimensional structure model is also encompassed within the scope of the present invention.

According to the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, one of the characteristics of the dipeptidyl peptidase IV can be found in that the dipeptidyl peptidase IV has 273 molecules of bond water in an asymmetric unit and has 5 molecules of N-acetylglucosamine residues per 1 molecule of the dipeptidyl peptidase IV.

According to the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, the information for atoms or atomic groups of the side chain of the dipeptidyl peptidase IV, interacting with the atoms or atomic groups of a known effector of the dipeptidyl peptidase IV via an intermolecular interaction can be obtained.

Furthermore, according to the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, the information of regions in the dipeptidyl peptidase IV that are susceptible to binding or intermolecular interaction with the effector can be obtained.

In addition, according to the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, the information of the structure specific to the dipeptidyl peptidase IV, which is not found in proteins other than the dipeptidyl peptidase IV, can be obtained. Therefore, higher selectivity in the effector targeting a protein other than the dipeptidyl peptidase IV can be designed, when the effector also acts on the dipeptidyl peptidase IV.

The intermolecular interaction includes covalent bond, ionic interaction, ion-dipole interaction, dipole-dipole interaction, hydrogen bonding, van der Waals force, electrostatic interaction, hydrophobic interaction and the like.

In the present specification, the atoms or atomic groups of the effector and  
5 atoms or atomic groups of the side chain of the dipeptidyl peptidase IV, which interact with each other via intermolecular interaction, are referred to as “pharmacophore.”

Also, according to the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, the information for the structure  
10 specific for the dipeptidyl peptidase IV, which is not found in proteins other than the dipeptidyl peptidase IV, can be provided.

In addition, for example, when the measurement conditions are different in X-ray diffraction, or the three-dimensional structure of the complex in the solution is analyzed using multidimensional NMR, and the like, the three-  
15 dimensional structural coordinate may differ from that shown in Figure 4. The three-dimensional structural coordinate varies depending on the fluctuation of protein and the like, and is encompassed within the scope of the present invention.

In the present specification, the “fluctuation of protein” means a state that  
20 is caused by molecular oscillation, temperature and the like, and accompanied with the structural change that can exhibit an activity for the dipeptidyl peptidase IV in a living body.

Also, according to the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, one of the characteristics of the  
25 dipeptidyl peptidase IV resides in that the amino acid residues, Ser 630, Asp 708

and His 740, which are involved in the activity deduced by experiments by using various active inhibitors of the dipeptidyl peptidase IV, exist in the adjacent area, even though the amino acid residues exist in distant locations on the primary sequence. Concretely, the distance between the  $O_{\delta 2}$  atom of Asp 708 and the  $N_{\delta 1}$  atom of His 740, and the distance between the  $N_{\epsilon 2}$  atom of His 740 and the  $O_{\gamma}$  atom of Ser 630 are distances that can form hydrogen bonding.

Therefore, the present invention also includes a three-dimensional structural coordinate of the region in the dipeptidyl peptidase IV, which is involved in binding or interaction of the dipeptidyl peptidase IV with an effector thereof, including a three-dimensional structural coordinate of a region selected from the group consisting of the following (a) to (d):

- (a) a region characterized by Ser 630, Asp 708 and His 740 in the amino acid sequence of SEQ ID NO: 2, and  
all or a part of a group of amino acid residues located in the adjacent area of each of the Ser 630, Asp 708 and His 740 in the structural coordinate shown in Figure 4 or the three-dimensional structure model defined by the structural coordinate;
- (b) a region characterized by Ser 630, Asp 708 and His 740 in the amino acid sequence of SEQ ID NO: 2, and  
all or a part of a group of amino acid residues comprising amino acids capable of maintaining physicochemical characteristics physiologically equivalent to each of amino acids of the group of amino acid residues located in the adjacent area of each of Ser 630, Asp 708 and His 740, in the structural coordinate shown in Figure 4 or the three-dimensional structure model defined by the structural coordinate,

(c) a region characterized by a group of amino acid residues comprising amino acids capable of maintaining physicochemical characteristics physiologically equivalent to each of Ser 630, Asp 708 and His 740 in the amino acid sequence of SEQ ID NO: 2, and

5 all or a part of a group of amino acid residues located in the adjacent area of said group of the amino acid residue in the structural coordinate shown in Figure 4 or the three-dimensional structure model defined by the structural coordinate; and

(d) a region characterized by a group of amino acid residues comprising amino acids capable of maintaining physicochemical characteristics physiologically equivalent to each of Ser 630, Asp 708 and His 740 in the amino acid sequence of SEQ ID NO: 2, and

10 all or a part of a group of amino acid residues of a group of amino acid residues comprising amino acids capable of maintaining physicochemical characteristics physiologically equivalent to the each amino acid of the amino acid residues located in the adjacent area of said groups of the amino acids, in the structural coordinate shown in Figure 4 or the three-dimensional structure model defined by the structural coordinate.

15 In the present specification, the "adjacent (area)" refers to an area involved in covalent bond, ionic interaction, ion-dipole interaction, dipole-dipole interaction, hydrogen bonding, van der Waals force, electrostatic interaction, hydrophobic interaction or the like with the amino acid residues, concretely, a region within 10Å, preferably within 8Å, more preferably within 5Å.

20 The physicochemical characteristic includes features in shape of the three-dimensional structure, hydrophobicity, electric charge, pK and the like.

The "amino acid capable of maintaining physicochemical characteristics physiologically equivalent" may be an amino acid analogue residue obtained by replacing a side chain of amino acid residues in the three-dimensional structural coordinate shown in Figure 4 with other side chain, for example, showing  
5 bioisosterism. Alternatively, the amino acid residue in the three-dimensional structural coordinate shown in Figure 4, may be replaced with another amino acid residue belonging to the same Group, in any of the following Groups I to VI:

- I glycine, alanine;
- 10 II valine, isoleucine, leucine;
- III aspartic acid, glutamic acid, asparagine, glutamine;
- IV serine, threonine;
- V lysine, arginine;
- VI phenylalanine, tyrosine.

15 According to the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, a three-dimensional structural coordinate of a polypeptide can be easily derived if an accurate amino acid sequence is determined, even when the polypeptide is a dipeptidyl peptidase IV or a dipeptidyl peptidase IV-like protein derived from other organism species, as long  
20 as the polypeptide is a polypeptide having high homology on the level of amino acid sequence with the dipeptidyl peptidase IV used for the preparation of the crystal of the present invention (for example, at least 20%, preferably 30% or more, more preferably 40% or more).

25 Furthermore, the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention can be used for X-ray crystallographic

structural analysis of the crystal and the like of other proteins having an amino acid sequence with significant homology with the dipeptidyl peptidase IV used for the preparation of the crystal of the present invention. Concretely, according to the molecular replacement method [for example, see Blundell, T. L. et al.,  
5 *PROTEIN CRYSTALLOGRAPHY*, 446-464 (1976), published by Academic Press and the like], the three-dimensional structural coordinate thereof can be quickly and readily obtained from the structure factors obtained by the X-ray diffraction pattern of the crystal, without using multiple isomorphous replacement method, even for the determination of the structural coordinate of the above-mentioned  
10 crystal of which structural coordinate has not yet been known.

In the present specification, the term "significant homology" is a case where there is identity of 20%, or more, preferably by 30% or more, between the amino acid sequences.

When the molecular replacement method is performed, for example, a  
15 program such as X-PLOR and CNX (both manufactured by Accelrys Inc.) or AMORE [one of the programs of CCP4 (Collaborative Computational Project, Number 4), *Acta Crystallogr. D*50, 670-673 (1994)] can be run by a computer on which the program can be executed. Here, whether or not the molecular replacement method is applicable can be determined by actually applying the  
20 molecular replacement method to the structure factors calculated from the X-ray diffraction pattern of the desired crystal and obtaining a significant solution.

In other words, the three-dimensional structural coordinate obtained by structural analysis by molecular replacement method is encompassed within the scope of the present invention as long as a significant solution is obtained. The  
25 present invention also encompasses a three-dimensional structural coordinate of

a dipeptidyl peptidase IV, or a dipeptidyl peptidase IV-like protein, namely a homolog protein or the like of other organism species derived by the above method.

Therefore, according to the present invention, a method for obtaining a three-dimensional structural coordinate of a homolog protein of a dipeptidyl peptidase IV comprising the step of performing refinement of an electron density map of the homolog protein of the dipeptidyl peptidase IV comprising the amino acid sequence of SEQ ID NO: 2, based on the three-dimensional structural coordinate of the present invention, to give a three-dimensional structural coordinate is provided. Also, a method for obtaining a three-dimensional structural coordinate of a crystal of a complex of a dipeptidyl peptidase IV and an effector of the dipeptidyl peptidase IV, based on the three-dimensional structural coordinate of the present invention, is likewise provided.

According to the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, a method for identifying a region or site for a target for binding or interaction between the dipeptidyl peptidase IV and an effector of the dipeptidyl peptidase IV is provided, based on the analysis of the binding regions between the dipeptidyl peptidase IV and a known effector of the dipeptidyl peptidase IV such as an inhibitor, or based on the simulation of the interaction between the dipeptidyl peptidase IV and a known effector of the dipeptidyl peptidase IV.

Also, based on the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention and the steric conformation of the effector of the dipeptidyl peptidase IV, the pharmacophore of the effector of the dipeptidyl peptidase IV can be identified. A method for identifying the

pharmacophore is also provided. The method is useful for designing an effector having excellent characteristics such as higher avidity, higher biological activity, higher biological stability, higher thermodynamic stability, higher absorbency to a living body, and lower toxicity.

5           Concretely, for example, the region or site for a target involved in binding or interaction of the dipeptidyl peptidase IV and an effector of the dipeptidyl peptidase IV, can be identified by:

1)     obtaining a crystal of a complex of the dipeptidyl peptidase IV and a known effector of the dipeptidyl peptidase IV such as an inhibitor, and obtaining  
10     a three-dimensional structural coordinate of the crystal based on the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention and the steric conformation of the effector of the dipeptidyl peptidase IV, whereby obtaining the three-dimensional structural coordinate of a binding region of the dipeptidyl peptidase IV and the effector;

15     2)     simulating an intermolecular interaction between the dipeptidyl peptidase IV and a known effector of the dipeptidyl peptidase IV based on the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention and the steric conformation of the effector of the dipeptidyl peptidase IV;

20     or the like.

The crystal of the above-mentioned complex can be obtained by, for example, incubating the crystal of the present invention in a solution comprising the effector, forming a complex of the dipeptidyl peptidase IV and the effector, and crystallizing the obtained complex, and the like.

25           Also, when the three-dimensional structural coordinate of the crystal of



the complex is obtained, the steric structure of the effector of the above-mentioned complex can be readily obtained by calculating the differential Fourier diagram utilizing a three-dimensional structure model defined by the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, whereby specific interaction forms and interaction sites  
5 between the dipeptidyl peptidase IV and the effector can be readily clarified.

When the intermolecular interaction is simulated, for example, the space regions, residues and the like in which covalent bond, ionic interaction, ion-dipole interaction, dipole-dipole interaction, hydrogen bonding, van der Waals  
10 force, electrostatic interaction, hydrophobic interaction or the like can be simulated, based on the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention and the steric conformation of the effector of the dipeptidyl peptidase IV.

Furthermore, according to the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, the three-dimensional  
15 structural coordinate or the three-dimensional structure model based on the three-dimensional structural coordinate regarded as an active center of the dipeptidyl peptidase IV, sites indirectly acting on the active center and regions or sites involved in binding or interaction with the effector, or the like, is obtained,  
20 whereby a compound capable of specifically acting on the dipeptidyl peptidase IV can be designed, identified, evaluated or searched.

For example, in the structural coordinate of Figure 4 and the three-dimensional structure model defined by the structural coordinate, a compound capable of modifying the activity of the dipeptidyl peptidase IV can be designed,  
25 identified, evaluated or searched, based on the regions characterized by Ser 630,

Asp 708 and His 740, and all or a part of amino acid residues of the group of the amino acid residues located in the adjacent area of the Ser 630, Asp 708 and His 740.

Therefore, according to the present invention, a method for designing,  
5 identifying, evaluating or searching an effector of the dipeptidyl peptidase IV is provided.

One of the significant features of the method of the present invention for designing, identifying, evaluating or searching an effector resides in that the method comprises designing, identifying, evaluating or searching a compound  
10 capable of acting on the dipeptidyl peptidase IV, based on the three-dimensional structural coordinate of the present invention.

According to the method of the present invention for designing, identifying, evaluating or searching an effector, since the method is based on the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, the information for a structure specific to the dipeptidyl  
15 peptidase IV, which is not found in proteins other than the dipeptidyl peptidase IV can be obtained. Therefore, according to the method of the present invention for designing, identifying, evaluating or searching an effector, the method has an excellent effect that the selectivity of the effector of the dipeptidyl peptidase IV  
20 can be enhanced.

Also, according to the method of the present invention for designing, identifying, evaluating or searching an effector, since the method is based on the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention, visual studies and/or energy calculation can be made  
25 according to the method by using a computer and the like. Therefore, there are

exhibited some excellent effects that a compound having excellent characteristics such as having higher avidity, higher biological activity, higher biological stability, higher thermodynamic stability, higher absorbency in a living body, and lower toxicity, than those for a known inhibitor can be designed, identified, evaluated or searched, and that logical design can be performed in the three-dimensional space.

In the present specification, the "effector" includes a compound that inhibits or enhances the activity (i.e. inhibitor or activator), which may be natural compounds or synthetic compounds, or may be polymers or low-molecular weight compounds.

A concrete example of the method of the present invention for designing, identifying, evaluating or searching an effector includes a method comprising the steps of:

- (i) identifying a region to be targeted for binding or interaction with the effector in a dipeptidyl peptidase IV, based on all and/or a part of the three-dimensional structural coordinate of the present invention and the steric conformation of the effector of the dipeptidyl peptidase IV;
- (ii) identifying corresponding atoms or atomic groups capable of generating in the region at least one intermolecular interaction selected from the group consisting of covalent bond, ionic interaction, ion-dipole interaction, dipole-dipole interaction, hydrogen bonding, van der Waals force, electrostatic interaction and hydrophobic interaction, with the atoms or atomic groups existing in a candidate compound; and
- (iii) designing a compound based on the above information of the above step (i) and/or (ii).

The three-dimensional structural coordinate used for designing, identifying, evaluating or searching a compound capable of binding to the dipeptidyl peptidase IV may be a coordinate fixed in the three-dimensional space, and the intensity of binding with the compound or the like can be calculated by carrying out translation or rotation in the three-dimensional space, and transfer to an extent that the chemical covalent bond would not be cleaved in the amino acid residues of the dipeptidyl peptidase IV.

In the above step (i), the "region to be targeted in the dipeptidyl peptidase IV" preferably includes an active center of the dipeptidyl peptidase IV, sites indirectly acting on the active center and the like. For example, there is included a region characterized by Ser 630, Asp 708 and His 740 and all or a part of a group of the amino acid residues located in the adjacent area of Ser 630, Asp 708 and His 740, and the like in the structural coordinate of Figure 4 and the three-dimensional structure model defined by the structural coordinate. The atoms or atomic groups that can be matched therewith are identified, based on the three-dimensional structural coordinate of an active center, sites indirectly acting on the active center and the like, whereby the candidate atoms or candidate atomic groups can be obtained.

In the above step (ii), for example, the atoms or atomic groups capable of associating via intermolecular interaction between the atoms or atomic groups in the region, concretely, the corresponding atoms or atomic groups capable of generating covalent bond, ionic interaction, ion-dipole interaction, dipole-dipole interaction, hydrogen bonding, van der Waals force, electrostatic interaction, hydrophobic interaction and the like, are searched and extracted, based on the information identified in the above step (i).

Next, in the above step (iii), the corresponding atoms or atomic groups searched in the above step (i) and/or (ii) are combined to design a compound.

Thereafter, if desired, whether or not the compound designed in the above step (iii) is matched via intermolecular interaction with the side chains and atoms or atomic groups in the dipeptidyl peptidase IV as defined by the three-dimensional structural coordinate of the present invention can be simulated.

The compound designed, identified, evaluated or searched by the above steps (hereinafter also referred to as a candidate compound in the present specification) can be obtained by generally used chemical synthetic methods, depending on the compound.

In addition, in the method of the present invention for designing, identifying, evaluating or searching an effector, there can be carried out a step of detecting the interaction between the dipeptidyl peptidase IV and the candidate compound. When the interaction is detected, the interaction serves as an index showing that the above candidate compound is a compound capable of binding to the dipeptidyl peptidase IV.

The above interaction can be detected by, for example, plasmon resonance analysis apparatus, mass spectrometer, titration isothermal calorimetry, NMR and the like. For example, in the case of plasmon resonance analysis apparatus, when a sensorgram indicates the formation of a complex, by contacting the dipeptidyl peptidase IV-immobilized matrix with the candidate compound and performing analysis by optical detection (for example, photometer, polarization photometer and the like) and the like, it would be an index showing that the interaction between the candidate compound and the dipeptidyl peptidase IV is generated. For example, in the case of a mass spectrometer, when a spectrum

indicates the formation of a complex, by contacting the dipeptidyl peptidase IV-immobilized matrix with the candidate compound and performing analysis with a mass spectrometer (matrix-assisted laser desorption/ionization-time of flight mass spectrometry: MALDI-TOF MS, electro spray-ionization mass

5 spectrometer: ESI-MS and the like), it would be an index showing that the interaction between the candidate compound and the dipeptidyl peptidase IV is generated. For example, in the case of titration-thermal calorimetry interaction analysis, when the titration curve indicates the formation of a complex, by contacting a solution of the dipeptidyl peptidase IV with the candidate compound, 10 and measuring the heat coming in and out of a thermal diode and the like, it would be an index showing that the interaction between the candidate compound and dipeptidyl peptidase IV is generated. For example, in the case of NMR, when a spectrum indicates the formation of a complex, by analyzing by NMR a solution prepared mixing the dipeptidyl peptidase IV and a candidate compound, 15 it would be an index showing that the interaction between the candidate compound and the dipeptidyl peptidase IV is generated.

Furthermore, the method of the present invention for designing, identifying, evaluating or searching an effector may further comprise the steps of contacting the dipeptidyl peptidase IV with a candidate compound, and thereafter 20 measuring the activity of the dipeptidyl peptidase IV. When the dipeptidyl peptidase IV activity increases or decreases, it would be an index showing that the candidate compound is a compound having enhancing action or inhibitory action on the activity of the dipeptidyl peptidase IV.

The dipeptidyl peptidase IV activity can be measured by, for example, 25 incubating a 1.5 ml reaction mixture [composition: 1.5 mM substrate

(Gly-Pro-paranitroanilide), 71 mM glycine-NaOH (pH 8.7)] at 37°C for 10 minutes in the presence of a candidate compound, and measuring the liberated paranitroanilide at the absorbance of 405 nm. During the measurement of the activity, the candidate compound may be evaluated by using a reaction system in which a suitable dilution series of the compound is added thereto.

The method of the present invention for designing, identifying, evaluating or searching the effector can be performed by, for example, sequentially selecting the interaction between the dipeptidyl peptidase IV and the compounds in a database in a computer to which the structures of plural of compounds had been inputted, or the interaction between the dipeptidyl peptidase IV and the designed compound, by visual methods (visual selection method) utilizing the database; and/or sequentially calculating the avidity with a computer, and searching a compound capable of stably interacting with the dipeptidyl peptidase IV from the database (computer-assisted avidity evaluation method) and the like, based on the three-dimensional structural coordinate of the present invention.

In the above visual selection method, the database of the structures of compounds may be a database in which the three-dimensional structural coordinates have been determined and inputted. Alternatively, in the case of a compound having a low molecular weight, the database may be a database in which the information for chemical covalent bond of a compound having a low molecular weight had been inputted, because the conformation can be relatively freely changed, and the three-dimensional structural coordinate of each conformation can be derived by calculation in a relatively short time.

Concretely, in the visual selection method, the expected complex between the dipeptidyl peptidase IV and a candidate compound or a part thereof is firstly

displayed on a computer screen, based on the three-dimensional structural coordinate of the present invention. Next, the intermolecular interaction binding between a compound in the database and the binding regions of the dipeptidyl peptidase IV is simulated on the computer, taking chemical interaction into consideration. Also, the simulation of the chemical modification of the compound is performed on the computer, and the changes in the interaction caused as a result thereof are observed on the computer screen. During the simulation, the three-dimensional space can be more easily understood by displaying the three-dimensional structure of the protein on the computer screen so that the structure corresponds to Crystal Eye glasses supplied by Silicone Graphics; simultaneously displaying two screens in which each angle is adjusted for displaying the object, according to the visual fields of the right eye and left eye, which is so-called referred to as "stereovision" which is frequently used by one of ordinary skill in the art; or the like. In addition, the three-dimensional structure can be visually studied by methods other than the stereoscopic displaying of the three-dimensional structure.

The candidate compound capable of generating suitable interaction can be obtained by displaying on a computer a group of candidates with appropriate conformation and selecting an appropriate one therefrom; calculating a structure having a low energy state on a computer; or the like. Next, a derivative of a compound capable of generating more preferable binding with the dipeptidyl peptidase IV may be searched among the candidate compound.

More specifically, on the level of the three-dimensional structure, the followings may be taken into consideration:

1 a group likely to be charged negatively, such as carboxyl group, nitro



group, or a halogen group in the compound interacts with an amino acid residue having a positive charge, such as lysine, arginine or histidine in the dipeptidyl peptidase IV;

2 a group likely to be charged positively, such as amino group, imino group  
5 or guanidyl group in the compound interacts with an amino acid residue having negative charge, such as glutamic acid or aspartic acid in the dipeptidyl peptidase IV;

3 a hydrophobic functional group such as an aliphatic group or an aromatic  
10 group in the compound interacts with a hydrophobic amino acid residue such as alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophane or methionine in the dipeptidyl peptidase IV;

4 a group involved in hydrogen bonding, such as hydroxyl group or amide  
group is allowed to form hydrogen bonding with a main chain or side chain portion;

15 5 a group or an atom likely to be charged negatively, such as carboxyl group, nitro group or a halogen group in the compound interacts with a positively charged atom on a main chain or side chain portion;

6 a group or an atom likely to be charged positively, such as amino group,  
20 imino group or guanidyl group in the compound interacts with a negative charged atom on a main chain or a side chain portion;

7 the flexibility of the three-dimensional structure of the compound is lowered by, for instance, cyclizing the linear chain portion;

or the like. For example, a derivative may be designed and synthesized so that the atoms having negative charge of the candidate compound are located in the

25 adjacent region of the side chain of an amino acid residue having positive charge

such as lysine, arginine or histidine, in the amino acid residue of the dipeptidyl  
peptidase IV, and that an atom having positive charge of the candidate  
compound is located in the adjacent region of the side chain of the amino acid  
residue having negative charge such as glutamic acid or aspartic acid in the  
5 amino acid residue of the dipeptidyl peptidase IV. Also, a group suitable for  
forming a hydrophobic interaction may be introduced into the portion capable of  
forming a hydrophobic interaction between the compound and the dipeptidyl  
peptidase IV, to design and synthesize a derivative. In addition, a group suitable  
for forming hydrogen bonding may be introduced into the portion capable of  
10 forming hydrogen bonding between the compound and the dipeptidyl peptidase  
IV, to design and synthesize a derivative. In the above-mentioned designing, it is  
desirable that van der Waals interaction is as high as possible, and that steric  
hindrance does not occur between each of the atoms. Furthermore, it is desirable  
that new void portions are not produced by modification of the compound and  
15 that in regions already containing void portions, the void portions are filled as  
much as possible.

As described above, the design, identification, evaluation or searching of a  
final compound can be thus performed with visually comprehensively  
considering intermolecular interaction and other factors on a computer screen.

20 In the computer-assisted avidity evaluation method, in order to determine  
the validity for the designing of a new compound, and to obtain a compound that  
can stably interact from the compounds in the database, a docking software  
(DOCK, GOLD, FlexX, Glide or the like) is used for evaluation of binding based  
on the energy by calculating a molecular force field between the compound and  
25 the dipeptidyl peptidase IV, evaluation of binding based on chemical

characteristics, evaluation of binding based on the Protein Data Bank (PDB), and the like. Further, in a model system consisting of the compound and the dipeptidyl peptidase IV, or in a model system further comprising solvent molecules and the like, it can be led to a compound that can stably interact by obtaining the index showing avidity, such as free energy of bonding, the ratio obtained from bond state number and non-bond state number, and the like by using molecular kinetic calculation or Monte Carlo calculation. The programs for calculation of molecular force field and molecular kinetic include AMBER, CHARMM, DISCOVER, PRESTO and the like, and the force field used includes AMBER, CHARMM, OPLS, MMCF, CVFF and the like. Furthermore, a program such as Ludi which automatically outputs the candidates for a candidate compound by providing a three-dimensional structural coordinate of the amino acid residues interacting in the dipeptidyl peptidase IV may be used.

The visual selection method and computer-assisted avidity evaluation method can be performed alone or in combination. In the case of performing the methods in combination, the avidity is actually calculated for the compounds that has been expected to be more desirable in visual investigation, and the validity thereof is evaluated. By repeatedly performing the calculation and evaluation, more excellent compounds may be designed, identified, evaluated or searched.

Next, the designed, identified, evaluated or searched compound is optimized to be a more excellent compound, such as a compound having more excellent characteristics as a medicament, such as being excellent *in vivo* kinetics, having low toxicity and low side-effect; a compound having a still higher biological activity as an effector; a compound having an advantageous structure as a medicament in view of its oral administration; and the like.

The resulting candidate compound can be obtained using generally used techniques for chemical synthesis depending on the kind of the compound.

The present invention also encompasses an effector of the dipeptidyl peptidase IV, which is obtained by the method of the present invention for designing, identifying, evaluating or searching an effector. When the effector is a compound capable of inhibiting or enhancing the activity of the dipeptidyl peptidase IV, the effector (inhibitor or activator) is expected to be an agent for, for example, a modulatory agent of immune response, a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like.

The three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention can be provided as a computer program, a medium or the like, which displays the three-dimensional structure of the molecule based on the three-dimensional structural coordinate and can be provided via a telecommunication line or the like. Therefore, using a computer or the like, the three-dimensional coordinate of the dipeptidyl peptidase IV can be displayed in detail, allowing to perform the method of the present invention for designing, identifying, evaluating or searching an effector more rapidly, conveniently and logically.

The present invention also encompasses a program or a medium therefor for use of the three-dimensional structural coordinate, in which all and/or a part of the three-dimensional structural coordinate of the dipeptidyl peptidase IV of the present invention is recorded.

The medium may be any of those in which the three-dimensional structural coordinate of the present invention can be derived on a program that

runs on a computer, and includes, for instance, electric memory media referred to as memory; semi-permanent memory media such as a FD, a hard disk, an optical disk, an opto-magnetic disk and a magnetic tape, and the like. In addition, the program and the medium therefor for use of the three-dimensional structural coordinate of the present invention also encompass those having a form which can be communicated via a telecommunication line such as internet.

Also, the program and the medium therefor for use of the three-dimensional structural coordinate of the present invention may further comprise a means for displaying the three-dimensional graphic display of the molecule. The program or the medium therefor which comprises the means for displaying the three-dimensional graphic display has advantages that visual studies and/or calculation of avidity can be made more conveniently, so that there is more facilitated a logical design on the three-dimensional structural level for obtaining a compound having excellent characteristics such as higher avidity, higher biological activity, higher biological stability, higher thermomechanical stability, higher absorbency to a living body, and lower toxicity than those for known effectors of the dipeptidyl peptidase IV.

As the means capable of displaying the three-dimensional graphic display, there may be generally used a program that is made so that a means for inputting the three-dimensional structural coordinate of the molecule, a means for measuring visual representation of the coordinate on a computer screen, the distance between the represented atoms in the molecule, bond angles or the like, a means for addition or modification of the coordinate, and the like can be provided. Furthermore, there may be used a program that has been made so that a means for calculating the structure energy of the molecule based on the

coordinate of the molecule, a means for calculating the free energy of bonding, and the ratio of bonding state number to non-bonding state number in consideration of solvent molecules such as water molecule can be provided.

Examples of the program suitable for such purposes include Insight II,

5 QUANTA and the like, which are computer programs commercially available from Accelrys Inc., and the present invention is not limited to these programs. Also, the above-mentioned programs can be introduced into a computer referred to as a work station supplied from Silicone Graphics Inc., SunMicro-Systems Ltd., or the like, and used.

10 According to the crystal of dipeptidyl peptidase IV of the present invention, there can be exhibited excellent effects that the three-dimensional structural coordinate can be obtained as an information for designing, identifying, evaluating or searching an effector of the dipeptidyl peptidase IV useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent

15 for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like, and that the crystal of a complex of the dipeptidyl peptidase IV and a known effector can be readily prepared.

Also, according to the three-dimensional structural coordinate of the present invention, there is exhibited an excellent effect that the effector can be designed,

20 identified, evaluated or searched. In addition, according to the method for obtaining a three-dimensional structural coordinate of the homolog protein of the dipeptidyl peptidase IV of the present invention, there is exhibited an excellent effect that the three-dimensional structural coordinate of the homolog protein of the dipeptidyl peptidase IV of which three-dimensional structure is unknown can

25 be conveniently and rapidly provided. Furthermore, according to the method for

obtaining a three-dimensional structure of a crystal of a complex of the dipeptidyl peptidase IV of the present invention and an effector of the dipeptidyl peptidase IV, there is exhibited an excellent effect that the method can provide a target for designing an effector useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like, and having excellent characteristics such as higher avidity, higher biological activity, higher biological stability, higher thermomechanical stability, and higher absorbency to a living body. Moreover, according to the method of the present invention for identifying a pharmacophore of the dipeptidyl peptidase IV and the effector of the dipeptidyl peptidase IV, there is exhibited an excellent effect that the method can provide a target for designing the effector useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like, and having excellent characteristics such as higher avidity, higher biological activity, higher biological stability, higher thermomechanical stability, and higher absorbency to a living body. According to the method of the present invention for designing, identifying, evaluating or searching an effector of the dipeptidyl peptidase IV, there is exhibited an excellent effect that the method can logically and conveniently provide an effector useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like, and having excellent characteristics such as higher avidity, higher biological activity, higher biological stability, higher thermomechanical stability,

and higher absorbency to a living body. According to the effector of the dipeptidyl peptidase IV of the present invention, there are exhibited excellent effects that the effector is capable of modifying immune response and capable of treating or preventing diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like. Furthermore, according to the program and medium therefor of the present invention, there is exhibited an excellent effect that the method for designing, identifying, evaluating or searching an effector of the dipeptidyl peptidase IV can be performed more rapidly and conveniently.

The present invention will be hereinafter more specifically explained by the following Examples, but the present invention is not intended to be limited by the Examples in any way. Unless otherwise indicated, the reaction conditions, procedures and the like can be referred to the instruction manual attached to the reagents used, *Molecular Cloning A Laboratory Manual*, third edition, Sambrook et al. [issued by Cold Spring Harbor Laboratory Press (2001)], and the like.

Example 1 Construction of Recombinant Baculovirus for Expression of Soluble Human Dipeptidyl Peptidase IV

(1) Cloning of Soluble Human Dipeptidyl Peptidase IV (shDPPIV) cDNA

Caco-2 cells [provided by American Type Culture Collection (ATCC)] were cultured at 37°C using Dulbecco's Modified Eagle Medium (manufactured by Invitrogen) containing 20% by volume of inactivated fetal bovine serum (manufactured by Invitrogen; inactivated by incubation at 56°C for 30 minutes) and 1% by volume of nonessential amino acid (manufactured by Invitrogen), in



the presence of 5% by volume of CO<sub>2</sub>.

Next, total RNA was extracted from the Caco-2 cells obtained.

Extraction of the total RNA was carried out using a product manufactured by Nippon Gene Co. Ltd. under the trade name of ISOGEN in accordance with the attached instruction manual. The obtained total RNA was used as a template for RT-nested PCR method described below.

In order to obtain a nucleic acid corresponding to a soluble human DPPIV from which the signal peptide sequence was removed (amino acid nos: 33-766 of SWISS-PROT Accession No: P27487), first, a cDNA fragment sequence of human DPPIV gene was amplified by RT-nested PCR method with total RNA of the Caco-2 as a template.

The thermal profile in the PCR is 30 cycles of reaction, in which one cycle comprises denaturation at 94°C for 30 seconds, annealing at 55°C for 30 seconds and polymerase extension reaction at 72°C for 1 minute.

The amplified DNA fragment was separated by agarose gel electrophoresis method, and a small fragment of the gel of the corresponding band portions was cut out. Thereafter, the DNA fragment was extracted from the obtained small fragments of the gel using a product manufactured by Bio 101 under the trade name of GENE CLEAN SPIN Kit, and purified. The obtained fragment was inserted into vector pCR2.1-TOPO contained in TOPO TA Cloning (registered trade mark) Kit manufactured by Invitrogen to construct pCR-shDPPIV.

In order to confirm whether or not the obtained cDNA fragment encodes the desired polypeptide, deletion mutants regarding the DNA fragment having various lengths were prepared, and a nucleotide sequence for the DNA fragment

was determined as follows.

First, a DNA fragment having a size of 2.2 kbp obtained by double digestion of the pCR-shDPPIV with *Bam*HI and *Eco*RI was inserted into a corresponding restriction site in pUC19 (manufactured by Takara Bio Inc.), to construct a plasmid pUshDPPIV. Various deletion plasmids were prepared using the plasmid pUshDPPIV by a conventional method.

The nucleotide sequence for the DNA fragment was determined using the obtained deletion plasmid or plasmid pCR-shDPPIV, and a product manufactured by Perkin-Elmer Cetus Inc. under the trade name of Taq DyeDeoxy Terminator Cycle Sequencing Kit and Model 373S sequencer manufactured by Applied Biosystems.

Also, the amino acid sequence of the polypeptide encoded by the above-mentioned DNA fragment was determined on the basis of the nucleotide sequence.

The determined amino acid sequence was compared with the sequence for a full length DPPIV of human colon shown in SEQ ID NO: 2. As a result, it was confirmed that the corresponding regions (regions excluding the transmembrane region) were identical.

Thus, it was confirmed that the DNA fragment encodes the desired polypeptide shDPPIV, namely a polypeptide in which the transmembrane region (amino acid nos: 1-32 at N-terminal side) in the full-length human DPPIV was deleted and a polyhistidine peptide was added to the C-terminal side.

## (2) Preparation of Recombinant Baculovirus

Plasmid pUshDPPIV was digested with a restriction enzyme to give a

DNA fragment encoding shDPPIV gene. The obtained fragment was inserted into pAcGP67B (manufactured by BD PharMingen) to construct a baculovirus transfer vector pAcGP67B-shDPPIV.

5 Fifteen minutes before the transfection, Sf21 cells were washed twice with a TNM-FH medium comprising 10% by volume of fetal bovine serum. The Sf21 cells were then transferred to a well of a 6-well plate by  $2.4 \times 10^6$  cells per well.

10 Furthermore, 2 to 5  $\mu\text{g}$  of the baculovirus transfer vector and a 0.5  $\mu\text{g}$  linear baculovirus DNA (trade name: BaculoGold virus DNA, manufactured by BD PharMingen) were mixed, and the mixture was allowed to stand at room temperature for 5 minutes. Next, 1 ml of Transfection Buffer B (manufactured by BD PharMingen) was added to the obtained mixture, and the mixture was thoroughly mixed to give a Transfection Buffer B/DNA mixture.

15 The medium in the wells of the 6-well plate and the cells that had not been adhered to the wells were removed, and 1 ml of Transfection Buffer A (manufactured by BD PharMingen) was added to each of the wells. The Transfection Buffer B/DNA mixture was gradually added dropwise to the wells of the 6-well plate, with gently stirring the 6-well plate. The cells were incubated at 28°C for 4 hours in the wells of the 6-well plate. Thereafter, the  
20 transfection buffer was removed, and 3 ml of TNM-FH medium containing 10% by volume of fetal bovine serum was added to the wells of the 6-well plate. The cells were cultured at 28°C in each of the wells of the 6-well plate for 5 days, and the culture supernatant was collected. The culture supernatant was used for  
25 amplification of virus using Sf21 cells to give a virus stock solution.

### Example 2 Preparation and Crystallization of shDPPIV

#### (1) Expression of shDPPIV in Insect Cells

Sf21 cells were cultured using a serum free medium EX-CELL 400 (manufactured by JRH Biosciences) and T flask, and Tn5 cells (provided by Invitrogen) were cultured using a serum free medium EX-CELL 401 (manufactured by JRH Biosciences) and a T flask, at 28°C, respectively. At the time when the proliferation of the cells reached 70% confluent, the old medium was removed, and a fresh medium was added at 40 ml per one 225-cm<sup>2</sup> flask. Then, 1.5 ml of virus stock solution after amplification for three times (having multiplicity of infection (MOI) of about 2) was added to the cells to infect the cells, and the cells were incubated at 28°C for 4 days. The culture supernatant four days after the infection was collected and stored at -20°C. The culture supernatant was used for the purification of shDPPIV protein as described below.

#### (2) Purification of shDPPIV Protein

In each step for the purification of shDPPIV, the activity of DPPIV was measured by incubating a 0.1 ml reaction mixture containing a 1.5 mM substrate [manufactured by Peptide Institute, Gly-Pro-paranitroanilide (pNA)], 71 mM Gly-NaOH (pH 8.7) and the DPPIV, and detecting the liberated pNA.

Meanwhile, the reaction mixture was incubated at 37°C for 10 minutes. During the incubation, the absorbance at 405 nm was monitored.

Also, the protein concentration was quantified by using a product manufactured by Bio-Rad Laboratories, Inc. under the trade name of DC protein Assay Kit II.

The purity of the protein was confirmed by subjecting a protein sample

in each step to SDS-PAGE using 7.5% polyacrylamide gel according to the method by Laemmli et al.

The culture supernatant stored at -20°C in the above-mentioned (1) was melted at 4°C and filtered with a bottle top filter (manufactured by Becton, Dickinson and Company) or with 0.45 µm filter (KURABO INDUSTRIES LTD.) to remove insoluble materials. The supernatant after the filtration was concentrated to an about tenth volume by using a concentrator Vivaflow 50 (manufactured by Sartorius AG) or Amicon stirrer cell model 8400 (manufactured by Millipore Corporation) to give a concentrated solution.

The obtained concentrated solution was dialyzed against buffer A (20 mM HEPES-NaOH, 0.5 M NaCl, pH 8.0) at 4°C overnight, and applied to a nickel column [one in which nickel was immobilized to HiTrap Chelating column (trade name, manufactured by Amersham-Pharmacia) (5 ml × 2)] equilibrated with buffer A. The column was washed with 10 column volumes of buffer A, and then with buffer A containing 50 mM imidazole. The elution of the fraction containing shDPPIV was carried out by a linear gradient of 50 to 500 mM imidazole. The fraction found to have DPPIV activity was collected, and dialyzed overnight at 4°C against buffer B (20 mM HEPES-NaOH, pH 8.0, 50 mM NaCl). After the dialysis, the sample was purified by using an anion exchange column [manufactured by Amersham-Pharmacia under the trade name: Resource Q (6 ml)] equilibrated with buffer B. The column was washed with buffer B, and thereafter shDPPIV was eluted by a linear gradient of 15 column volumes of 50 to 500 mM NaCl. The fractions found to have DPPIV activity were collected, and used as a purified preparation.

### (3) Preparation of Protein Sample for Crystallization

The shDPPIV purification sample (9 ml) obtained in the above (2) was concentrated using a product manufactured by Millipore Corporation under the trade name of Centricon 10 until the protein concentration reached 10 mg/ml.

5 The obtained product was used as a protein sample for crystallization. The protein sample for crystallization was stored at 4°C.

A precipitation agent solution containing 0.18 M glycine-NaOH (pH 9.5), 0.18 M sodium sulfate and 18% by weight of PEG4000, and a 10 mg/ml dipeptidyl peptidase IV solution were mixed, and thereafter, a drop of the  
10 obtained mixed solution was placed on a product under the trade name of Cryschem Plate (manufactured by Hampton Research). The above-mentioned precipitation solution was allowed to stand at 20°C as a reservoir solution to allow crystallization.

### 15 (4) Crystallization of shDPPIV

The crystallization of shDPPIV was carried out by a sitting-drop method, which is one of vapor diffusion methods.

The formation of crystal was observed with the passage of time using a stereoscopic microscope. As a result, after about two weeks, a large crystal  
20 having a maximum size of 500  $\mu\text{m}$   $\times$  300  $\mu\text{m}$   $\times$  100  $\mu\text{m}$  was obtained. The crystal is also referred to as a native crystal. The microphotograph of the obtained crystal is shown in Figure 1. In Figure 1, the visual field is 4000  $\mu\text{m}$   $\times$  3000  $\mu\text{m}$ .

### 25 Example 3 Three-Dimensional Structural Analysis of Crystals

(1) X-ray Diffraction

The crystal obtained in Example 2 mentioned above was soaked in a cryoprotecting buffer [composition: 0.18 M glycine-NaOH (pH 9.5), 19% by weight of PEG4000, 0.18 M sodium acetate, 15% glycerol], and immediately thereafter the mixture was placed under nitrogen gas stream (100 K) to rapidly freeze the mixture.

The X-ray diffraction intensity data of the above crystal were collected up to the resolution of  $3.0\text{\AA}$  using a product manufactured by Rigaku International Corporation under the trade name of R-AXIS IV in nitrogen gas stream (100 K), and converted to the structure factor using a program MOSFLM (Version 6.11). A photograph of the diffraction pattern is shown in Figure 2.

From the obtained diffraction intensity data, it was determined that the crystal form to which the crystal belongs was orthorhombic, that the space group was  $P2_12_12_1$ , and the lattice constants were  $a = 118.0 \pm 5.0\text{\AA}$ ,  $|b| = 125.9 \pm 5.0\text{\AA}$  and  $|c| = 136.8 \pm 5.0\text{\AA}$ .

(2) Multiple Isomorphous Replacement Method

In order to derive an electron density map, multiple isomorphous replacement method was carried out. The crystal obtained in Example 2 was soaked for 3 days and 4 days in a crystallization solution prepared by dissolving mercury chloride until being saturated, to give two different kinds of isomorphous replacement crystals containing mercury atoms in the crystals. The X-ray diffraction intensity data were collected in the same manner as those for the native crystals.

In the determination of the phase in the structural analysis, CCP4

(Collaborative Computational Project, Number 4, 1994. "The CCP4 Suite: Programs for Protein Crystallography," *Acta Cryst.* D50, 760-763) program was used.

5 First, Fourier transform calculation utilizing the difference between the diffraction intensity obtained from the two kinds of isomorphous replacement crystals of mercury and the diffraction intensity obtained from the native crystal was performed using MLPHERE contained in the CCP program package. The position of each mercury atom in the unit cell of the real space was determined by investigating large peaks provided by heavy atoms (mercury) in the obtained  
10 Patterson's diagram. The phase of the crystal structure factor of the native crystals was determined by using the obtained position coordinate of mercury atoms. Furthermore, in order to determine the coordinate of each mercury atom more accurately using DM and SOLOMON contained in the CCP program package, refinement was carried out using three crystal structure factors of the  
15 native crystals and of the two kinds of mercury isomorphous replacement crystals.

An electron density map of the crystals of the dipeptidyl peptidase IV in real space was obtained using the phase of the crystal structure factor of the native crystals calculated from the refined coordinates of the mercury atoms.  
20 Furthermore, the electron density map was improved by carrying out smoothening and histogram matching of the electron density map in a solvent region, to obtain an electron density map critical for molecular modeling.

### (3) Molecular Modeling

25 The sites corresponding to the amino acid residues of the dipeptidyl



peptidase IV were identified on the electron density map by using QUANTA (manufactured by Accelrys, Inc.), to build molecular models.

As expected from the lattice constants, there were two molecules of the dipeptidyl peptidase IV in an asymmetry unit, and a model was built for each of the molecules. The refinement of the obtained molecular model was carried out using CNX (manufactured by Accelrys, Inc.), and the molecular model was adjusted again using the QUANTA for the obtained improved electron density map. The procedures were repeated to build a more accurate molecular model. In the refinement of the final coordinate, diffraction intensity data measured again were used after OSMIC confocal mirror (manufactured by Rigaku International Corporation) had been introduced into R-Axis IV (trade name, manufactured by Rigaku International Corporation).

As a result, the resolution was improved from the previous 3.0Å to 2.6Å. Furthermore, 273 molecules of bound water and 5 molecules of N-acetyl glucosamine residues per molecule of the dipeptidyl peptidase IV were identified in an asymmetric unit. R factor, which is an index for accuracy of the obtained molecular model, was 24.89%, and a free R factor, which independently was not taken into account of the calculation of refinement at the step of refinement, was 30.15%. During the procedure, the deviation of the interatomic bond distance (rms-deviation) and the bond angle from the ideal state of the three-dimensional structure were 0.006Å and 1.305°, respectively. The stereogram of the three-dimensional structure model of the crystals is shown in Figure 3, and the coordinate is shown in Figure 4. The present coordinate data were registered in PDB (Brookhaven Protein Data Bank) [PDB Code No: 1J2E, RSCB code No: 005544].

Here, as to those regions which did not take a regular structure in the crystals (in the disordered state), namely, the region from Asp 38 to that closer to the N-terminal side thereof, and the region for the tagged peptide (polyhistidine peptide) of the C-terminal side, the molecular model could not be built.

5 Furthermore, a part of the side chains projected to the surface of the molecules did not take a regular structure. However, these residues were not portions that play an important role for the function of enzymes.

In the three-dimensional structure of the dipeptidyl peptidase IV, which has been clarified by the Examples, it has been revealed that the amino acid  
10 residue involved in the activity deduced by various experiments for the dipeptidyl peptidase IV, namely, Ser 630, Asp 708 and His 740, form hydrogen bonds between the O<sub>δ2</sub> atom of Asp 708 and N<sub>δ1</sub> atom of His 740, and with the N<sub>ε2</sub> atom of His 740 and O<sub>γ</sub> atom of Ser 630, even the residues locate in distant locations on the primary sequence. Therefore, for the structural coordinate of  
15 Figure 4 and the three-dimensional structure model defined by the structural coordinate, it is suggested that the regions characterized by Ser 630, Asp 708 and His 740, and the whole or a part of amino acid residues that are located in the vicinity of Ser 630, Asp 708 and His 740 play an important role on the exhibition of the activity for the dipeptidyl peptidase IV and binding or interaction of the  
20 dipeptidyl peptidase IV with the effector, and that the compound matching the three-dimensional structure of the regions affect the activity for the dipeptidyl peptidase IV.

The present invention may be embodied in other various forms without  
25 departing from the spirit or essential characteristics thereof. The present

embodiment is therefore to be considered in all respects as illustrative and not restrictive, the scope of the invention being indicated by the appended claims rather than by the foregoing description and all changes which come within the meaning and range of equivalency of the claims are therefore intended to be embraced therein.

### INDUSTRIAL APPLICABILITY

According to the crystal of the dipeptidyl peptidase IV of the present invention, the information of a three-dimensional structure coordinate suitable for designing, identifying, evaluating or searching an effector of the dipeptidyl peptidase IV, useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like can be obtained. Also, according to the three-dimensional structure coordinate, the information of a three-dimensional structure coordinate suitable for designing, identifying, evaluating or searching an effector of the dipeptidyl peptidase IV, useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like can be obtained. Further, according to the method of the present invention for obtaining a three-dimensional structure coordinate of a homolog protein of a dipeptidyl peptidase IV, the refinement of the three-dimensional structure coordinate of the homolog protein of the dipeptidyl peptidase IV can be more conveniently carried out. Moreover, according to the method of the present invention for obtaining a three-dimensional structure coordinate of a crystal of a complex of a dipeptidyl

peptidase IV with an effector of the dipeptidyl peptidase IV, the information of a three-dimensional structure coordinate suitable for designing, identifying, evaluating or searching an effector of the dipeptidyl peptidase IV, which is useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like, and is excellent in avidity, biological activity, biological stability, absorbency to a living body, and which can favorably act on the dipeptidyl peptidase IV can be obtained. Also, according to the method for identifying a pharmacophore of a dipeptidyl peptidase IV and an effector of the dipeptidyl peptidase, the information of a three-dimensional structure coordinate suitable for designing, identifying, evaluating or searching an effector of the dipeptidyl peptidase IV, which is useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like, and is excellent in avidity, biological activity, biological stability, absorbency to a living body, and which can favorably act on the dipeptidyl peptidase IV can be obtained. Further, according to the method of the present invention for designing, identifying, evaluating or searching an effector of the dipeptidyl peptidase IV, the information of a three-dimensional structure coordinate suitable for designing, identifying, evaluating or searching an effector of the dipeptidyl peptidase IV, which is useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like, and is excellent in avidity, biological activity, biological stability, absorbency to a living body,

and which can favorably act on the dipeptidyl peptidase IV can be logically and conveniently obtained. In addition, the effector of the dipeptidyl peptidase IV of the present invention is useful as a modulatory agent of immune response and as a therapeutic or prophylactic agent for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like. Further, according to the program or the medium therefor of the present invention, the design, identification, evaluation and search for an effector of a dipeptidyl peptidase IV can be carried out rapidly and conveniently. Therefore, the present invention can be utilized in modulation of immune response and the treatment or prevention for diabetes, inflammation, multiple sclerosis, Graves' disease, chronic rheumatoid arthritis, AIDS, cancer and the like.